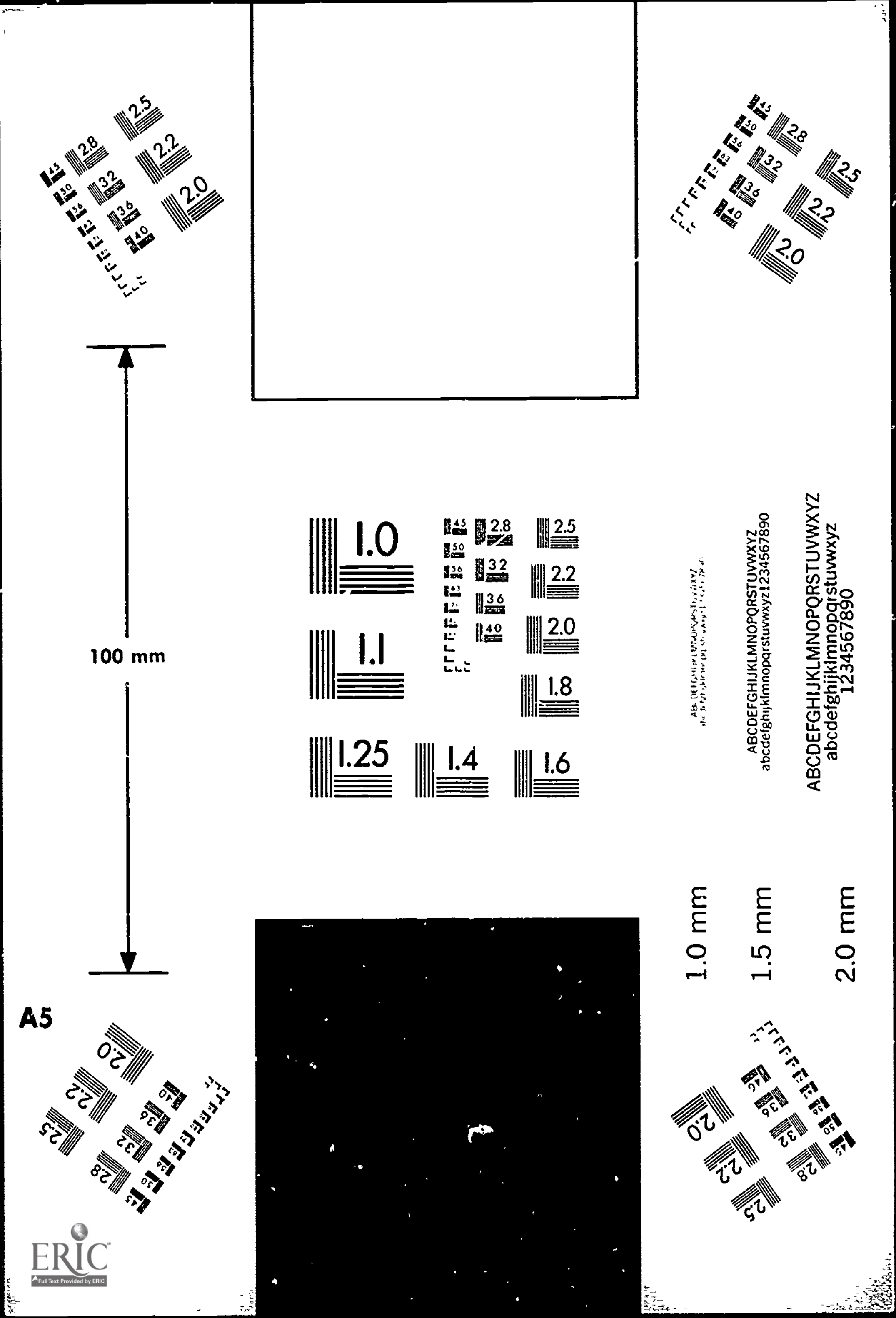

100 mm
A5
1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz1234567890
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz
1234567890
1.0 mm
1.5 mm
2.0 mm
ERIC
Full Text Provided by ERIC

DOCUMENT RESUME

ED 320 368 EC 231 182

AUTHOR	Farlow, Leslie J.; And Others
TITLE	The Implications of the Procedural Contrast between Training and Probe Conditions on the Interpretation of Student Performance Data.
SPONS AGENCY	Office of Special Education and Rehabilitative Services (ED), Washington, DC.
PUB DATE	Mar 88
GRANT	G008530150
NOTE	21p.; Paper presented at the Annual Convention of the Council for Exceptional Children (66th, Washington, D.C., March 28-April 1, 1988).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Behavioral Science Research; Daily Living Skills; *Data Interpretation; Elementary Secondary Education; *Instructional Effectiveness; *Performance Factors; *Research Methodology; *Severe Disabilities; Student Evaluation; *Training Methods

ABSTRACT

The study examined the relationship between probe and training conditions on performance of nine severely handicapped students on functional tasks in classroom and community settings. Findings led to the following recommendations to teachers: measures of student performance under natural or probe conditions should be evaluated in addition to training data; teachers should use different criteria for evaluating multiple-opportunity probes than for single-opportunity probes; teachers should not assume training data are representative of performance under natural conditions; and, teachers should not assume that performance under training conditions will be depressed since, in some cases, performance under probe conditions may be higher than performance under training conditions. Contains 20 references. Tables provide statistical detail. (DB)

The Implications of the Procedural Contrast Between Training and Probe Conditions on the Interpretation of Student Performance Data

Leslie J. Farlow

Indiana University at South Bend

Brenda H. Loyd

and Martha E. Snell

University of Virginia

Paper presented at the national conference of The Council for Exceptional Children, Washington, D.C., March-April, 1988

[1]This research was supported by Grant G008530150, Office of Special Education and Rehabilitation Services, U.S. Department of Education. The authors gratefully acknowledge Gail Munger for her assistance in analysis of data.

Running Heading: INTERPRETATION OF CONTRAST

The Implications of the Procedural Contrast
Between Training and Probe Conditions
on the Interpretation of Student Performance Data

Special education teachers monitor students' learning and instruction by collecting data on student performance. These data are often evaluated through visual analysis of charts (Snell, 1988; White & Haring, 1980). However, the appearance of the data may be influenced by the conditions in effect during data collection. Teachers may collect data under instructional conditions (with training procedures implemented) or under probe conditions (with no reinforcement or assistance available to the student). The conditions in effect during data collection may have an impact on student performance and therefore would have an impact on teacher interpretation of data. To date, there has been little research on the impact of this procedural contrast to guide teachers in data interpretation.

Data collected under probe conditions can be especially useful to teachers. These data provide teachers with information about student performance under natural conditions, about the types of errors they make, and about adaptation and generalization (Horner, Sprague & Wilcox, 1982; Liberty, 1985, Snell & Grigg, 1987; Wilcox & Bellamy, 1982). However, probe conditions may negatively influence student performance. Cuvo (1979) argued that students may react to probes as they would to extinction. The withdrawal of assistance and reinforcement may be distressing to students.

Research about this phenomenon of procedural contrast has been inconclusive. Buchwald (1959a, 1959b, 1960) conducted a series of studies on the effect of feedback on college students. The students were trained to select syllables from word cards with various combinations of reinforcement and

feedback. In a total of 5 experiments, the students returned to baseline levels of performance after repeated exposure to probe conditions (i.e., no reinforcement or feedback). Duker and Morshink (1984) found that four students with profound mental retardation, who had been trained to use manual signs, stopped using the signs in settings where the signs were not reinforced, another example of extinction.

In contrast to the above experiments, Frankel, Simmons, Fichter, and Freeman (1984) and Schriebman, Koegel, and Craig (1977) found students with autism and mental retardation actually improved performance under probe conditions with repeated exposure to the probe conditions. In these studies students were taught to respond to multiple cues (e.g., auditory and visual) using reinforcement. Then students were probed on responses to one cue without reinforcement. In both studies, students' performance improved after exposure to probes. Farlow, Snell, and Loyd (1987) analyzed the relationship between probe and training data from classrooms for students with severe handicaps and found students sometimes actually performed higher on probe trials than on training trials.

The previous research has other limitations which restricts its interpretation. Results based on Buchwald's work (1959a, 1959b, 1960) with college students may not be generalizable to students with severe handicaps. All of the studies described above involved non-functional tasks in laboratory settings (Buckwald, 1959a, 1959b, 1960; Frankel, et al., 1984; Schriebman, et al., 1977). When subjects used signs to make requests, their requests were ignored which made the use of the requests nonfunctional (Duker & Morsink, 1984). Farlow, et al. (1987) used a statistical rather than visual analysis to interpret data and the results of this study were confounded by the inclusion of more students with high rates of maladaptive behaviors in one comparison group.

The current study further explores the relationship between probe and training conditions on student performance of functional tasks, in classroom and community settings. Statistical procedures were employed, but results were compared with a visual analysis of the data. Five hypotheses were addressed. The first two hypotheses examined the relationship between probe and training data. If training data reflect student learning under natural conditions, then probe and training scores would be related. If students react to probe conditions as they react to extinction, then probe performance would be lower than training performance.

Additional hypotheses addressed the type of probe procedures teachers use: (a) single-opportunity (probe is ended after first error), or (b) multiple-opportunity (student is given the opportunity to perform all of the steps on the task). Snell and Browder (1986) in a review of literature on community-referenced instruction found both types of probe procedures were used. If the type of probe condition effects student performance, comparisons and interpretation of research could be effected. Grigg, Snell, and Loyd (1989) found that teachers of students with severe disabilities did consider the type of probe when evaluating student performance data.

Three hypotheses about the difference between probe procedures were tested. First, because the student is given the opportunity to perform all of the steps in the task, it was expected that performance on multiple-opportunity probes would be more closely related to training performance than performance on single opportunity probes. Second, because the student is given the opportunity to perform all the steps on a task under the multiple-opportunity probe condition, it was expected that the difference between probe and training scores would be less under the multiple opportunity condition than under the single-opportunity probe condition. Third, because the student is stopped after the first error under the single opportunity condition, performance may be more likely to

fluctuate because of errors on the first few steps, therefore, it was expected that performance on single-opportunity probes were more likely to be variable than performance on multiple opportunity probes.

Method

Sample

Thirteen teachers in the central Virginia region submitted over 500 records of student performance (including program formats and student performance data) for 54 students with severe handicaps. The program data reported in this study were selected at random from those 500 records. Student records also had to meet two criteria: (a) a student had at least one instructional program for which single opportunity probes were used and one for which multiple-opportunity probes were used to measure student performance; (b) both training and probe data had been collected for a minimum of 25 trials on each program.

Data were analyzed for 18 programs of nine students. All of the students had severe to profound retardation, four had additional physical or sensory impairments, and four students had high rates of maladaptive behaviors (i.e., teachers reported using contingencies for maladaptive behaviors for all or almost all of the students' instructional programs). The skills included were all functional and represented domestic, vocational, leisure, and community domains. Programs addressed the skills of: greeting, signaling for communication, self-care, assembly tasks, playing a tape recorder, shopping and using a restaurant. Instructional strategies included system of least prompts (11 programs), time delay (6 programs), and graduated guidance (1 program).

Procedure

Each probe data point was paired with the training point that occurred closest in time to compute the statistics for measuring each of the five

hypotheses. Data from the first 15 training sessions were not included so that the student had an opportunity to distinguish between probe conditions and training conditions. Ten to thirty pairs of points for each program were used for analysis. The number of pairs varied according to the amount of time of program implementation and the ratio of probe trials to training trials.

To determine if training data reflected student performance under natural conditions, Pearson product moment correlations were computed for each pair. To determine if students reacted to probe conditions in the same way they react to extinction, the difference between each probe point and the temporally closest training point was computed.

The repeated measures T-test (SPSSX, 1986) was computed to examine the three hypotheses concerned with difference between probe procedures. To determine if multiple-opportunity probe performance was more like training performance than performance on singe-opportunity probes, a T-test was computed on the differences in student performance under probe and training conditions and the difference in correlations for the multiple-opportunity probe and single-opportunity probe groups. Finally, to determine if student performance was more likely to be variable under the single-opportunity probe condition than under the multiple opportunity-condition, the difference between standard deviations according to probe condition was tested.

A post-hoc analysis was conducted to determine if the probe condition groups differed on program characteristics. The difference in the number of steps on the task analysis and the number of steps entered for each program were measured by a T-test. Differences between teachers, skill domain area, and instructional strategies were tested using a chi-square procedure.

Results

Reliability

Reliability data were available for 11 of 18 programs included in this analysis. Data were collected on 2 to 37% of the trials for each program. Reliability per session range from 60 to 100%. Average reliability per program ranged from 89% to 100%. In all cases reliability data were collected by an observer who recorded student performance independent of the teacher. The role for the independent observer was varied. In some instances, the teacher collected reliability data when classroom assistants ran the program; other reliability data were collected by graduate students enrolled in a teacher education program, and some reliability data were collected by research assistants for this study.

Data analyzed in this study were from teachers who had received advanced training in the education of students with severe to profound disabilities. Their training had included data collection skills. All teachers collected data regularly on all instructional objectives. Finally, the school districts employing the teachers supported consistent data collection procedures. Therefore, the reliability data which was reported was assumed to be representative of these teachers' data collection practices.

Individual Program Comparisons

The results of individual program analyses are presented in Table 1. Correlations for instructional programs that employed multiple-opportunity probes ranged from -.03 to .83. Five of the nine correlations were significant at the .05 level. Under the single-opportunity probe condition, correlations ranged from .06 to .47, and two of the nine correlations were significant at the .05 level.

Insert Table 1 about here

Comparison of Probe Conditions

Results of the difference between probe condition groups is presented in Table 2. There was no statistical difference between probe conditions for correlation with training, or for training variance. Variance of student performance was greater under the single-opportunity probe condition; and the difference between scores on training and probe are greater when using single-opportunity probes.

The post-hoc analysis for difference between groups is illustrated in Table 3. There was no statistical difference between groups based on the number of steps in the task analysis, the number of data pairs entered for analysis, the domain, or the instructional strategy. There was a significant difference between groups based on the teacher who implemented the program.

Insert Tables 2 and 3 about here

Comparison to Visual Analysis

Student performance data are typically evaluated through visual analysis. Statistical analysis of data may not yeild the same results as visual analysis (Center, Skiba, & Casey, 1985; Furlong & Wampold, 1981; Huitema, 1986; Jones, Weinrott, & Vaught, 1978). Data for the programs included in the analysis were also charted for visual analysis.

If a teacher evaluated the trend of probe performance and the trend of training performance to be similar, the performance under training and probe conditions would be correlated. Those programs that were significatly correlated did show

similar trends. Examples of charts for significantly correlated and not significantly correlated programs are illustrated in Figure 1.

The degree of mean difference between probe and training data should be indicated by differences in the level of charted data. Those programs with larger mean differences scores also appeared different in visual analysis, (i.e., the level of performance under probe conditions appeared to be higher or lower than performance under training conditions). These differences are also illustrated in Figure 1.

Place Figure 1 about here

Standard deviations were used to present the variability of data for the statistical analysis. Standard deviation measures the distance from the mean for each point. Data showing an accelerating trend may have had standard deviations similar to data that were variable and flat. The 10 graphs with the highest standard deviations (range from 21 to 46) were evaluated with visual analysis. Of the 10 cases, 9 did show an accelerating trend. However, of those, 8 were also variable. Variable, for the purpose of visual analysis, was defined as more than 80% of the points falling more than 20% away from the trend line (Browder, 1987).

Discussion

This study addressed five questions about the relationship between probe and training conditions. The results of correlational analyses did not show a consistent relationship between probe and training data. Further analyses indicated that the relationship did not seem to be influenced or explained by: a) the type of probe condition, b) the type of instructional procedure, c) the length of the task analysis, or d) the domain area of the skill. These results

suggest that a teacher who is concerned about student performance under probe conditions is not assured of making the same judgements if evaluating training data.

The next question addressed was whether probe conditions would effect performance in ways similar to extinction. The mean difference scores indicate that students do not perform consistently lower on probe conditions. In this study, students performed better under probe conditions than under training conditions in half of the cases. The conclusion reached by Buchwald (1959a; 1959b; 1960) and Duker and Morsink (1984), that students react to probe conditions in the same ways they react to extinction conditions, was not supported by this research.

The assumptions about multiple and single-opportunity probes were also addressed. This study supported the assumptions that: a) single-opportunity probes were more likely to appear variable than multiple-opportunity probes, and b) scores on single-opportunity probes were more likely to differ from training scores than will scores on multiple-opportunity probes. This information may influence a teacher's evaluation of single-opportunity probe data. For example, a teacher might be less concerned about variability of probe data on single-opportunity probes during acquisition phases of a program.

The results of this study replicated the results of the earlier Farlow et al. (1986) study. In the earlier study, the results were confounded due to a difference in student characteristics in the two probe condition groups. This study included the same students in both groups. This study also compared statistical analysis of data with visual analysis of the same data. The relationships indicated through correlations and standard deviations were similar to relationships observed through visual analysis.

This research suggests several guidelines for teachers who are evaluating student data. Measures of student performance under natural or probe conditions should be evaluated in addition to training data. Teachers should use different criteria for evaluating multiple-opportunity probes than for single-opportunity probes. Teachers should not assume training data are representative of performance under natural conditions. Finally, teachers should not assume that performance under training conditions will be depressed. In some cases, performance under probe conditions may be higher than performance under training conditions.

References

Browder, D.M. (1987). Assessment of individuals with severe handicaps: An applied behavior approach to life skills assessment. Baltimore: Brookes.

Buchwald, A. M. (1959a). Experimental alterations in the effectiveness of verbal reinforcement combinations. Journal of Experimental Psychology, 57, 351-361.

Buchwald, A. M. (1959b). Extinction after acquistion under different verbal reinforcement combinations. Journal of experimental Psychology, 57, 42-48.

Buchwald, A. M. (1960). Supplementary report: Alternation of the reinforcement value of a positive reinforcer. Journal of Experimental Psycholgy, 60, 416-418.

Center, B.A., Skiba, R.J., Casey, A. (1985). A methodology for the quantitative synthesis of intra-subject design research. Journal of Special Education, 19, 387-400.

Cuvo, A. J. (1979). Multiple-baseline design in instructional research: Pitfalls of measurement and procedural advantage. American Jouranl of Mental Deficiency, 84, 219-288.

Duker, P. C., & Morsink, H. (1984). Acquisition and cross-setting generalization of manual signs with severely retarded individuals. Journal of Applied Behavior Analysis, 17, 93-103.

Farlow, L.J., Loyd, B.K., & Snell, M.E. (October, 1986). Assessing Student Performance: The effect of procedural contrast between training and probe data. Paper presented at the annual meeting of the Association for Persons with Severe Handicaps.

Frankel, F., Simmons, J. Q., Fichter, M., & Freeman, B. J. (1984). Stimulus and overselectivity in autistic and mentally retarded children. Journal of Child Psycholgy and Psychiatry, 25, 147-155.

Furlong, M.J., & Wampold, B.E. (1981). Intervention effects and relative variations as dimensions in experts' use of visual inference. Journal of Applied Behavior Analysis, 15, 415-421.

Horner, R.H., Sprague, J. & Wilcox, B. (1982). General case programming for community activities. In B. Wilcox and G.T. Bellamy (Eds.), Design of high school programs for severely handicapped students (pp. 61-98). Baltimore: Paul H. Brookes.

Huitema, B.E. (1986). Autocorrelation in behavioral research: Wherefore art thou? In A. Poling and R.W. Fuqua (Eds.), Research methods in applied behavior analysis: Issues and advances (pp. 187-208) New York: Plenum Press.

Jones, R.R., Weinrott, M.R., & Vaught, R.S. (1978). Effects of serial dependency on the agreement between visual and statistical inference. Journal of Applied Behavior Analysis, 11, 277-283.

Shriebman, L., Keogel, R. L., & Craig, M. S. (1977). Reducing stimulus overselectivity in autistic children. Journal of Child Pshchology, 5, 425-436.

Snell, M.E. (1987). Systematic instruction of persons with severe handicaps (3rd ed.). Columbus, OH: Charles E. Merrill Publishing Company.

Snell, M. E., & Grigg, N. C. (1987). Instructional assessment and curriculum development. In M. E. Snell (Ed.), Systematic instruction of persons with severe handicaps (3rd ed.), (pp. 64-109). Columbus, OH: Charles E. Merrill Publishing Company.

Snell, M.E. & Browder, D.M. (1986). Community-referenced instruction: Research and issues. Journal for the Association of Persons with Severe Handicaps, 11, 1-11.

SPSSX (1986). Chicago: SPSS, Inc.

White, O., & Haring, N. (1980). Exceptional teaching. Columbus, OH: Charles E. Merrill.

Wilcox, B., & Bellamy, G.T. (1982). Instructional methods. In B. Wilcox and G.T. Bellamy (Eds.), Design of high school programs for severely handicapped students (pp. 99-120). Baltimore: Paul H. Brookes.

Table 1

Individual Program Analysis

	Multiple Opportunity						Single Opportunity					
	Probe		Training				Probe		Training			
Ss	$\bar{X}$	SD	$\bar{X}$	SD	D	r	$\bar{X}$	SD	$\bar{X}$	SD	D	r
1	61	19	65	17	4	.83***	52	29	85	14	33	.44*
2	65	11	64	11	-1	.36*	31	13	73	11	42	.35
3	70	11	69	12	-1	.03	67	33	84	23	17	.33
4	32	15	35	11	3	.43*	79	30	85	22	6	.06
5	83	19	81	19	-2	.59***	59	32	72	34	13	-.08
6	92	9	88	8	-4	.16	84	21	94	9	10	.47
7	74	13	79	17	5	.60	67	37	49	46	-18	.12
8	64	18	62	21	-2	.42	63	30	54	41	- 9	.30
9	58	34	52	32	6	.43*	26	15	24	18	- 2	.47**

*$p \leq .05$. **$p \leq .01$. ***$p < .001$.

Table 2

Comparison of Probe Conditions

	Mean	SD	T-Value
Correlation Between Probe & Training			
Multiple-opportunity	.42	.25	1.26
Single-opportunity	.27	.20	
Probe Variance			
Multiple-opportunity	16.55	7.56	-2.45*
Single-opportunity	26.82	8.30	
Training Variance			
Multiple-opportunity	16.49	7.10	-1.78
Single-opportunity	24.34	13.45	
Difference Between Probe & Training			
Multiple-opportunity	3.01	1.72	-2.92**
Single-opportunity	16.54	13.07	

*$p < .05$. **$p < .01$.

Table 3

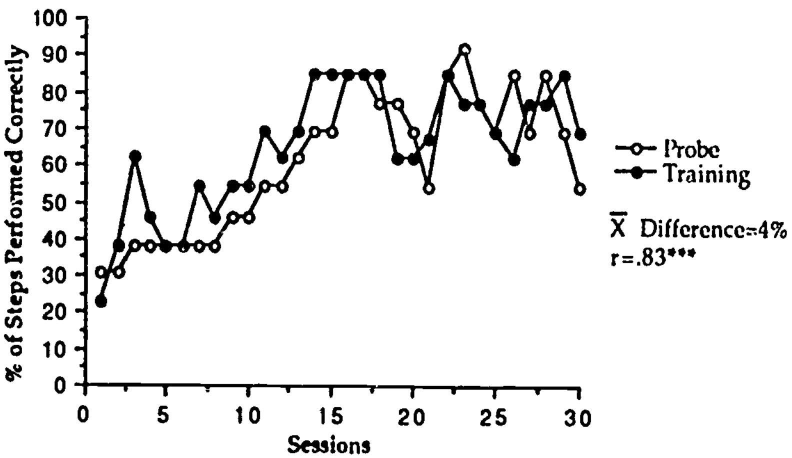
Results of Test for Equivalence of Groups

Characteristic	T-Value
Steps in the Task Analysis	-1.46
Pairs entered	.91
	X^2
Teacher	9.00*
Domain	6.00
Instructional Strategy	2.91

*p < .05.

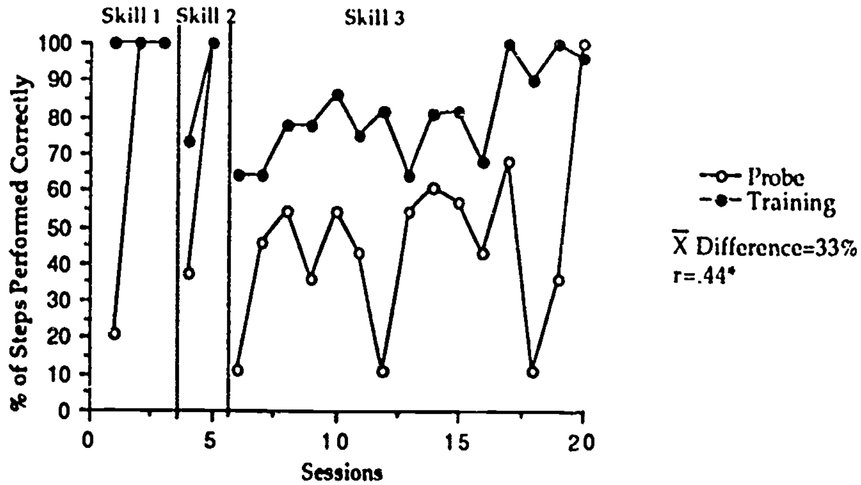
Figure 1. Examples of charted student performance data illustrating mean difference and high and low correlations between probe and training performance.

High Correlation, Low Mean Difference



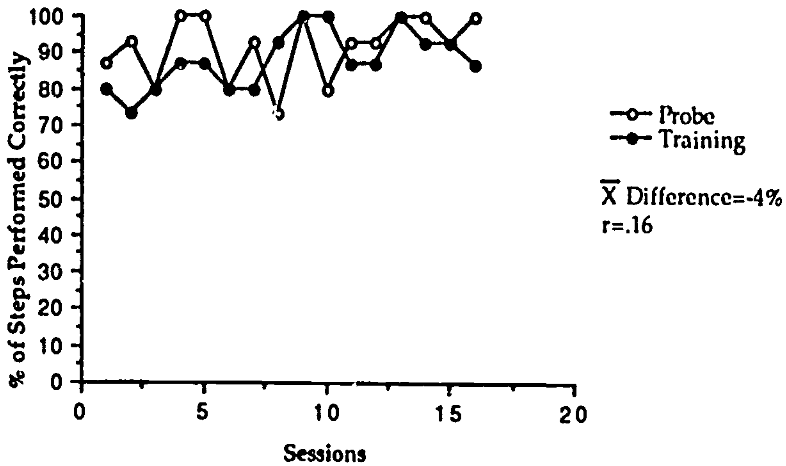
Student #1, Wash Hands
Multiple Opportunity Probe

High Correlation, High Mean Difference



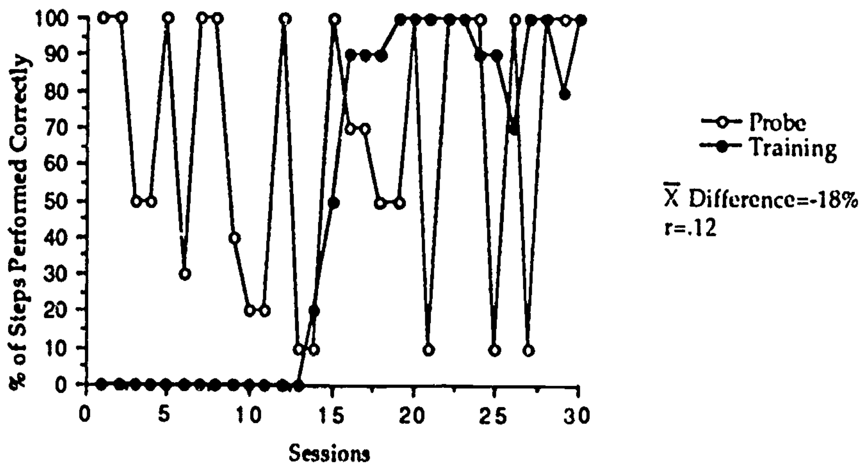
Student #1, Use Convenience Store (3 skills)
Single Opportunity Probe

Low Correlation, Low Mean Difference



Student #6, Play Tape Recorder
Multiple Opportunity Probe

Low Correlation, High Mean Difference



Student #7, Use Fast Food Restaurant
Single Opportunity Probe